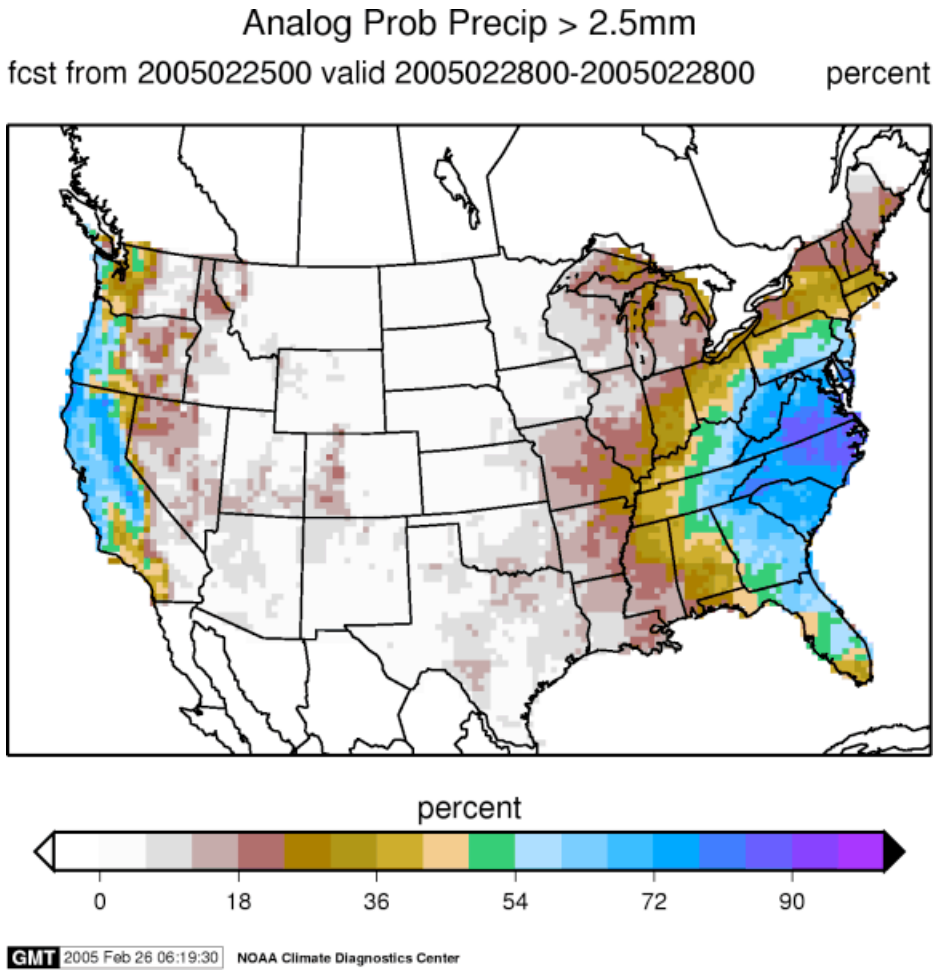
New Experimental PQPF Products from CDC

Jeff Whitaker, Tom Hamill and Xue Wei
NOAA Climate Diagnostics Center, Boulder
jeffrey.s.whitaker@noaa.gov

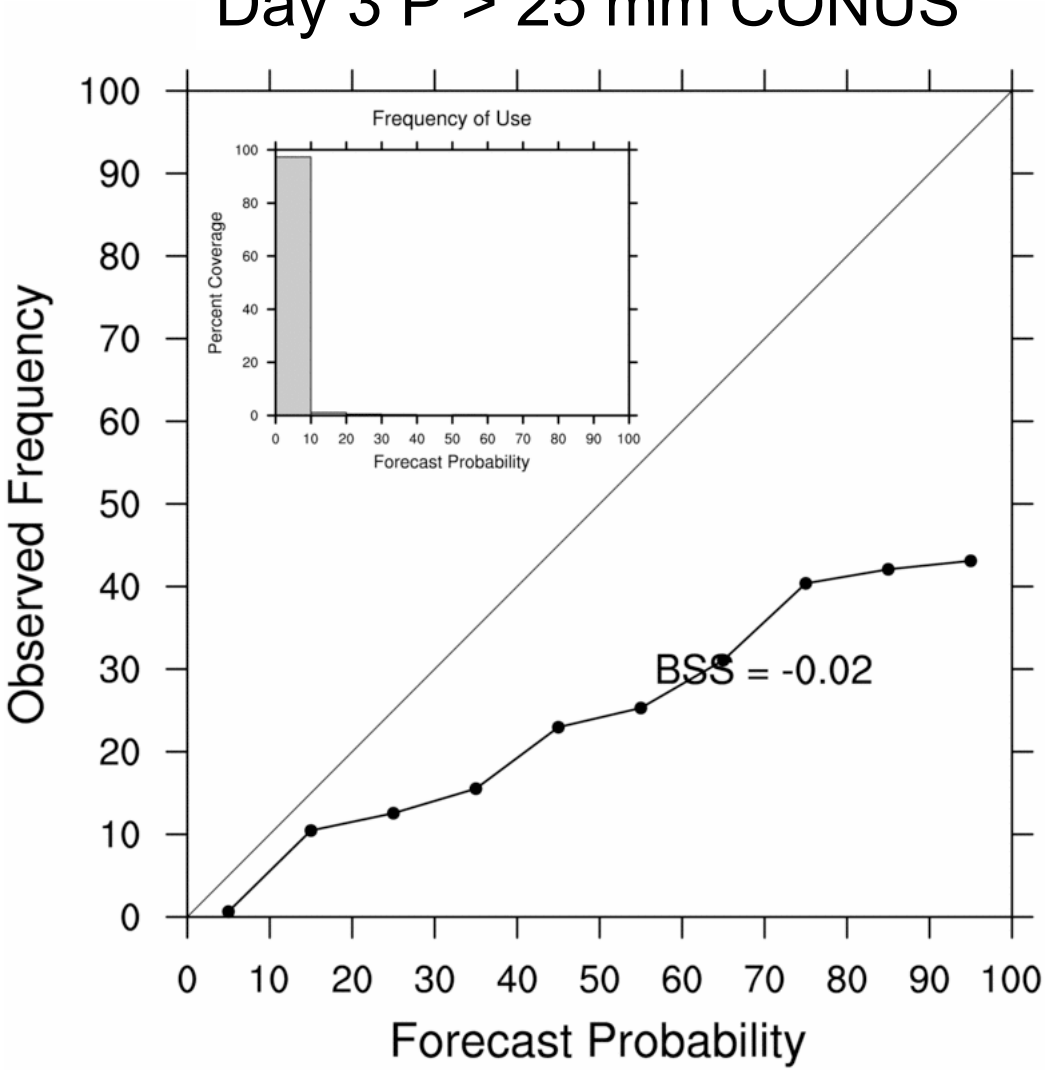


Why PQPF is hard

- Large model systematic errors (in mean and shape of probability distribution - e.g. the drizzle problem).
- Lots of spatial detail not resolved by models.
- Raw probabilities from ensembles have low skill.

Skill of GFS EPS (JFM 2002/2003)

Day 3 P > 25 mm CONUS



What do we want?

- PQPF is a conditional probability - the probability distribution for observed weather given a forecast from a model.
- Using ensemble relative frequencies assumes that the forecast PDF is the same as the conditional probability of observed weather (model is perfect).

How do we get it?

Conditional climatology of observed wx given a model forecast.

- **Forecast Analogs**: Given today's forecast, find prior cases when forecast was close to it, use corresponding analyses of observed weather.
- **Problem**: Need to wait a REALLY long time to find good analogs! (Lorenz, van den Dool).
- **Solution**: Search for analogs over a limited region.

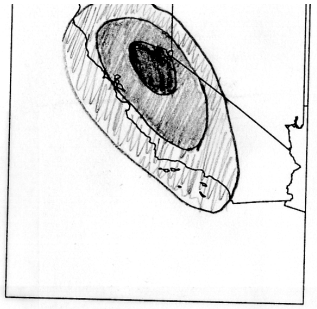
Analog technique:

(pioneered by van den Dool, Toth, von Storch, others)

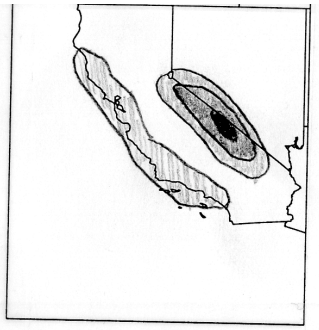
Step 2: find dates of old analogs

Step 3: extract observed weather

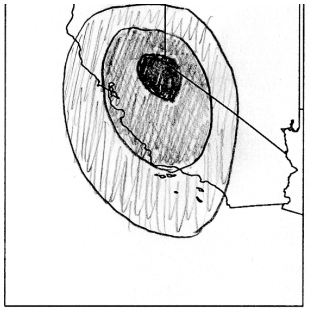
Forecast analog 1, 2/12/95



Observed Wx, 2/12/95

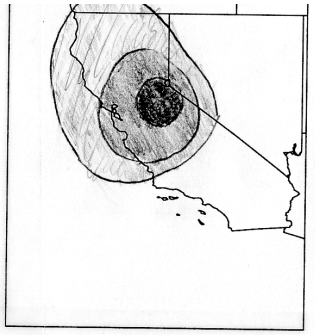


TODAY' S MODEL PRECIP FORECAST

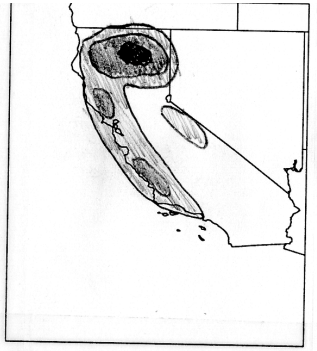


Step 1: make today's forecast

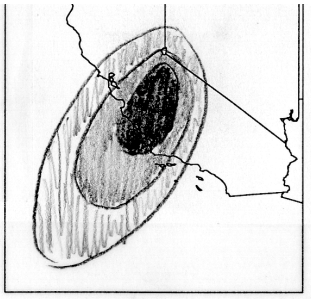
Forecast analog 2, 1/16/98



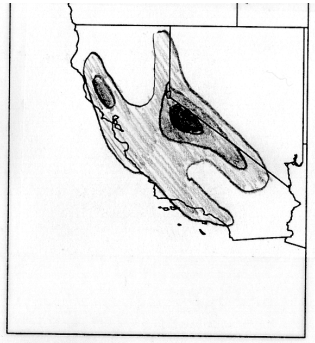
Observed Wx, 1/16/98



Forecast Analog 3, 3/1/83

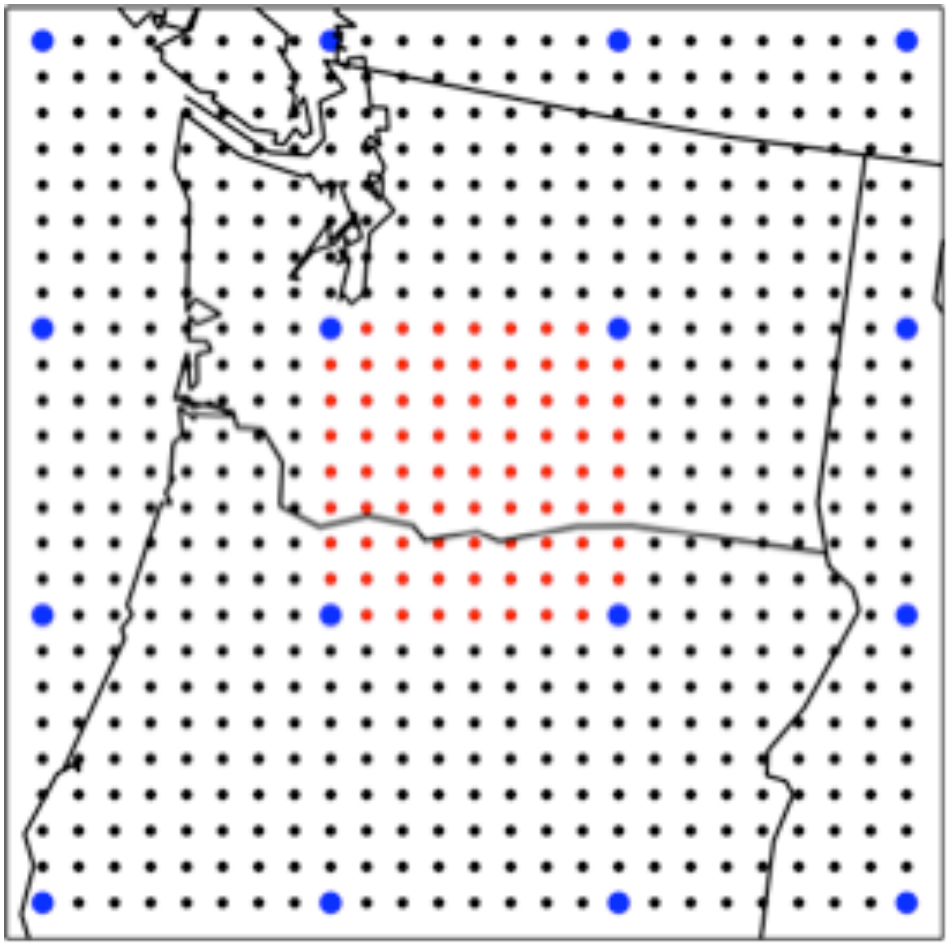


Observed Wx, 3/1/83



How analog ensemble is created

- analogs are found in for global model forecasts at blue points
- high-resolution NARR precip for those dates at red points make up analog forecast ensemble.
- process is repeated for next overlapping 16x16 box, forecasts are stitched together



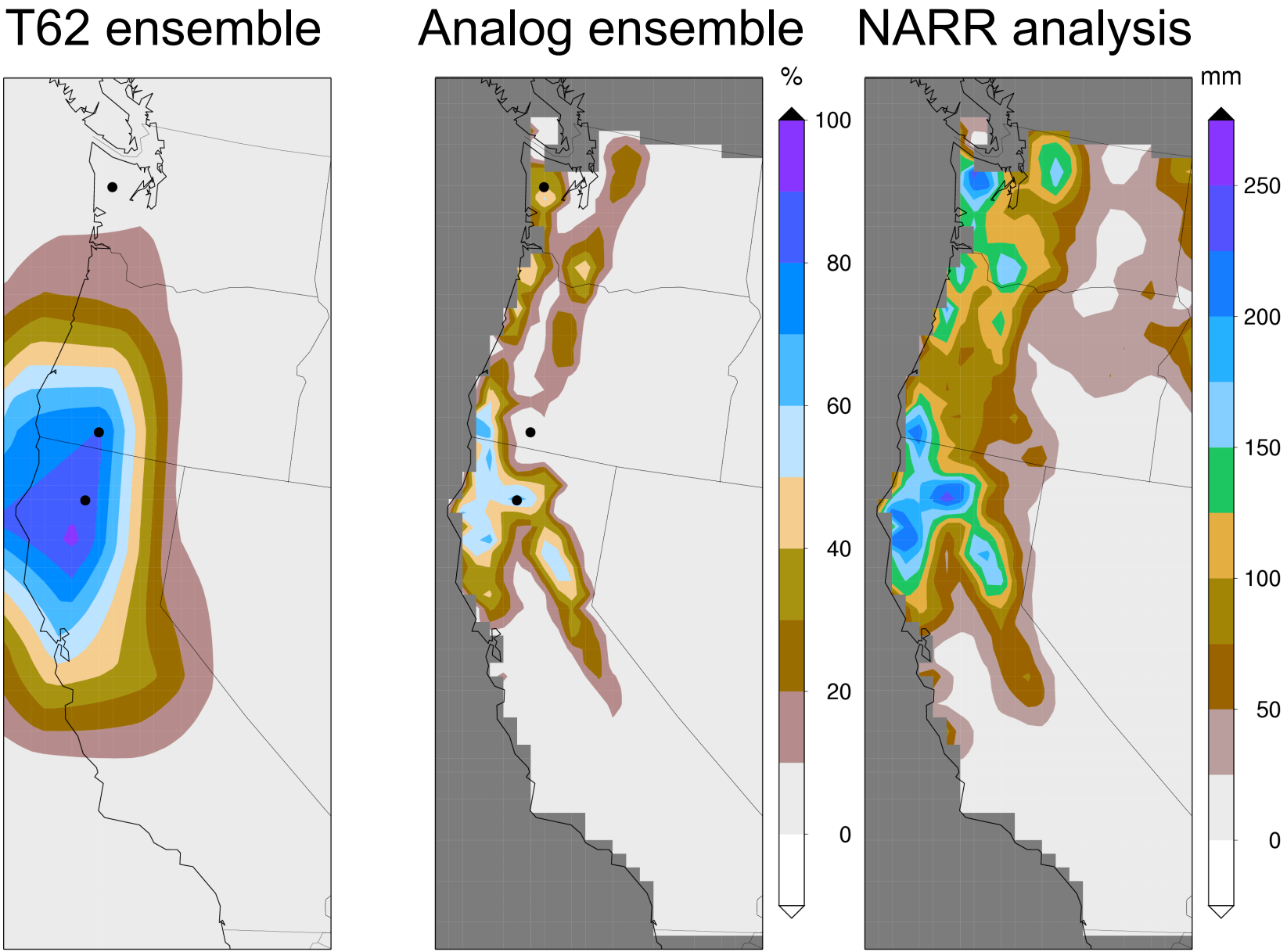
CDC MRF reforecast data set

- **Definition of "reforecast"** : a data set of retrospective numerical forecasts using the same model to generate real-time forecasts.
- **Model**: T62L28 NCEP MRF, circa 1998 (*http://www.cdc.noaa.gov/people/jeffrey.s.whitaker/refcst* for details).
- **Initial states**: NCEP-NCAR reanalysis plus 7 +/- bred modes (Toth and Kalnay 1993).
- **Duration**: 15-day runs every day at 00Z from 19781101 to now. (*http://www.cdc.noaa.gov/people/jeffrey.s.whitaker/refcst/week2).*
- **Data**: Selected fields (winds, geo ht, temp on 5 press levels, and precip, t2m, u10m, v10m, pwat, prmsl, rh700). (Web form to download at *http://www.cdc.noaa.gov/reforecast*).
- **Verifications**: North American Regional Reanalysis (NARR) - 32km grid.

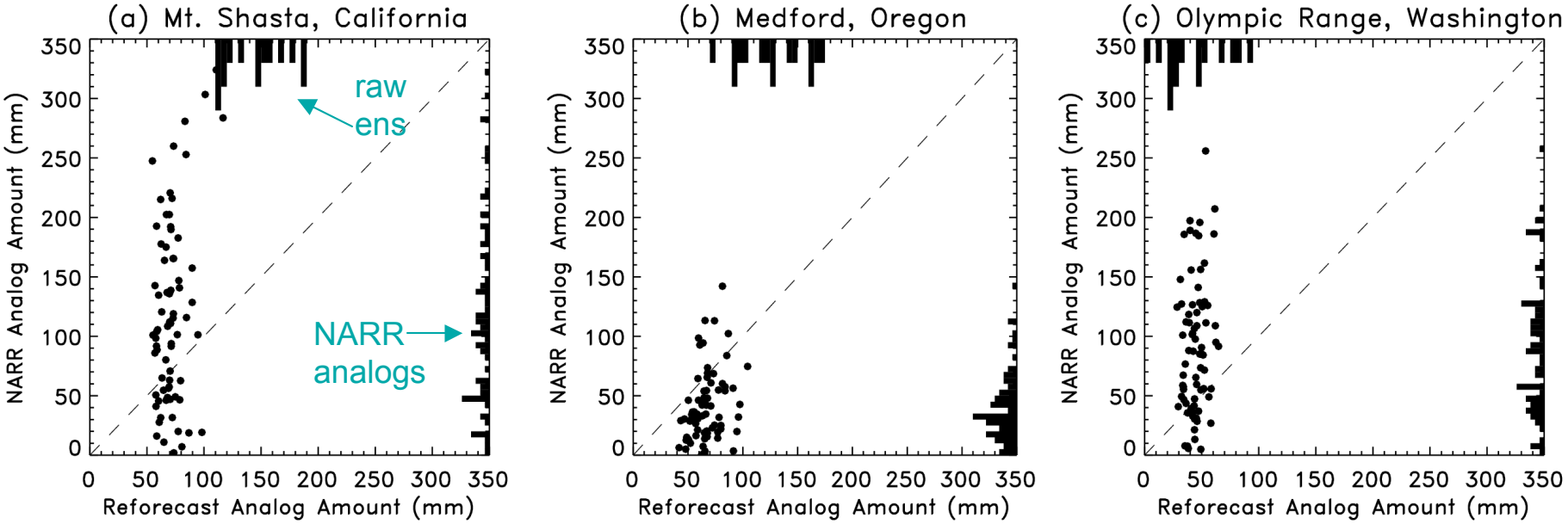
Positives/Negatives

- **(+)** If you have LOTS of data, can't do any better (no approximation).
- **(+)** Non-parametric, can be applied to any quantity as long as obs are available.
- **(-)** Need LOTS of data to find good analogs (decades, not centuries for local analogs).
- **(-)** Can't handle unprecedented events.

Example: P > 100mm, days 4-6 (19971229-19970101)



What's happening?

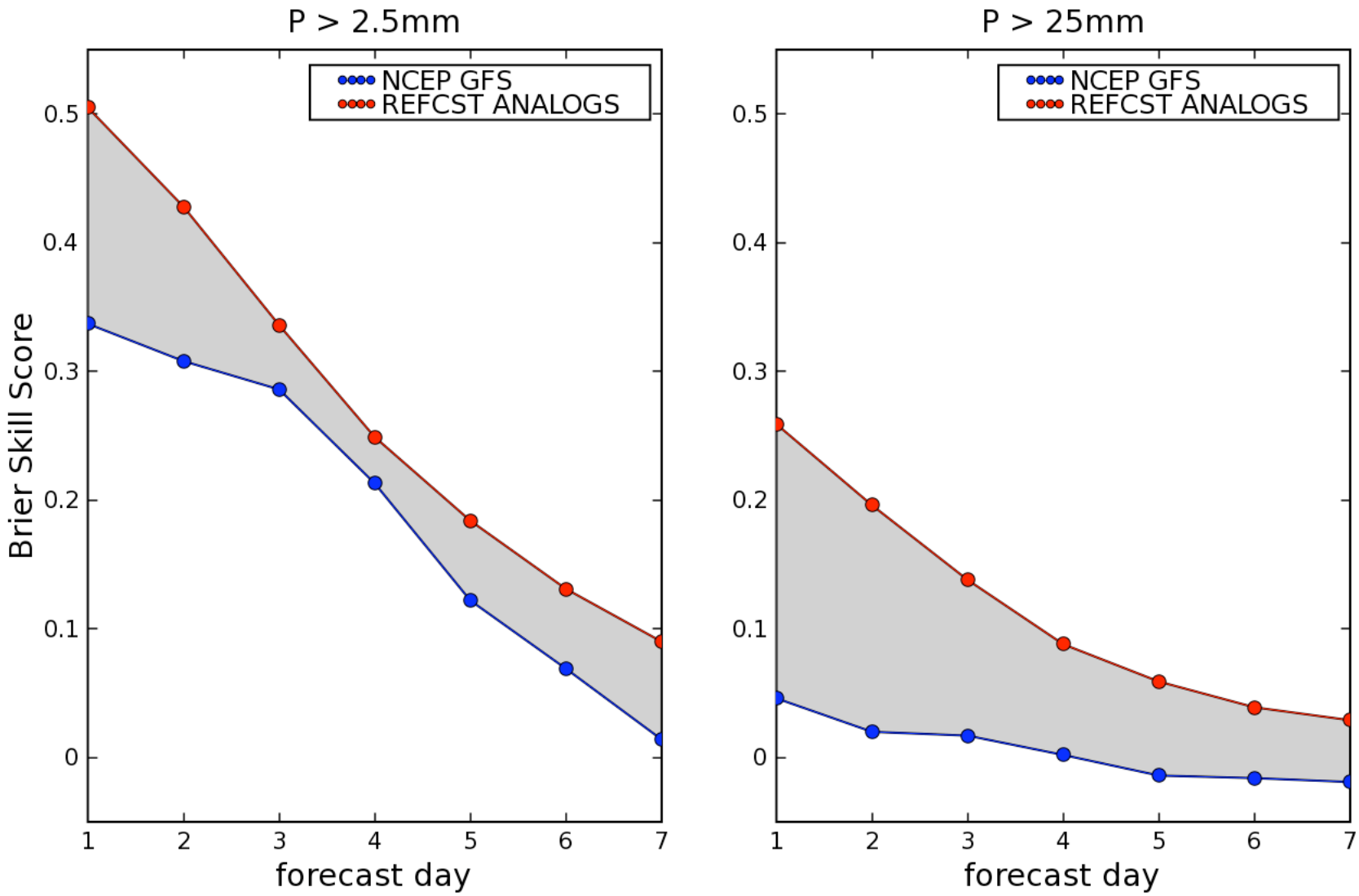


Can't find any other reforecast analogs with precip as heavy. But introduce large scatter by taking associated observed analogs. (P reduced)

Again, few close reforecast analogs. But observed data recognizes overforecast bias. (P reduced)

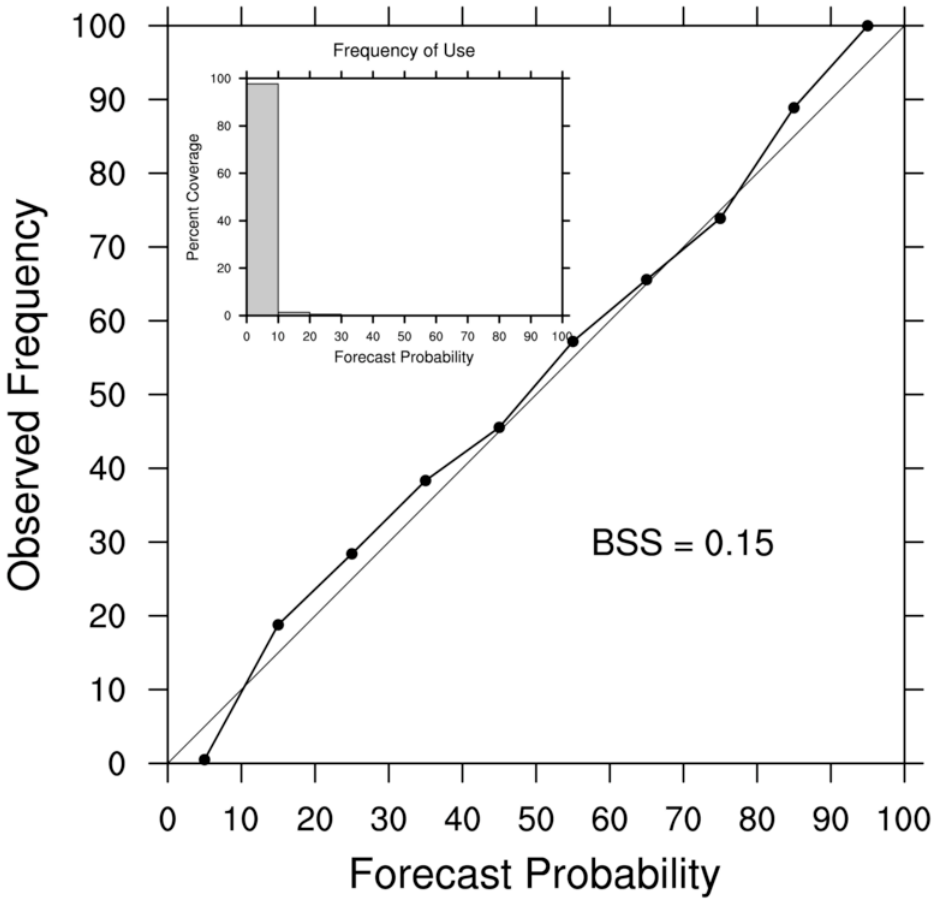
Here there are close reforecast analogs. Observed data introduces spread, increases amount. (P increased)

Skill comparison

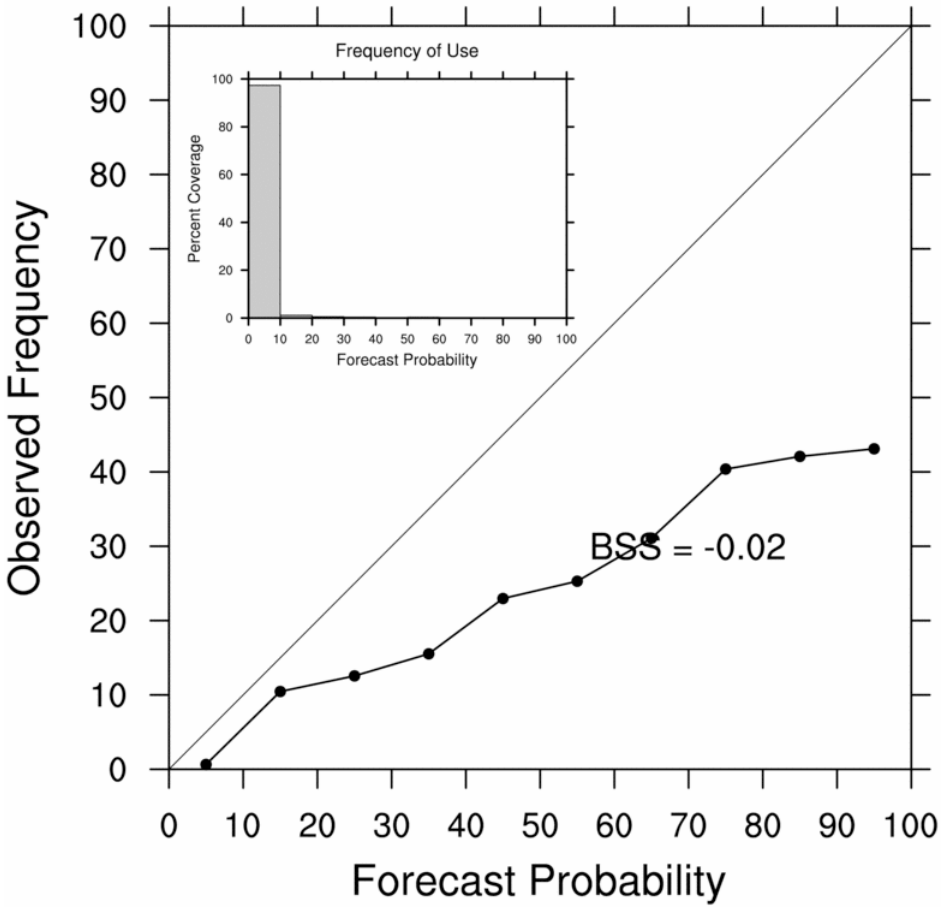


Day 3 P > 25mm, JFM 2002/2003

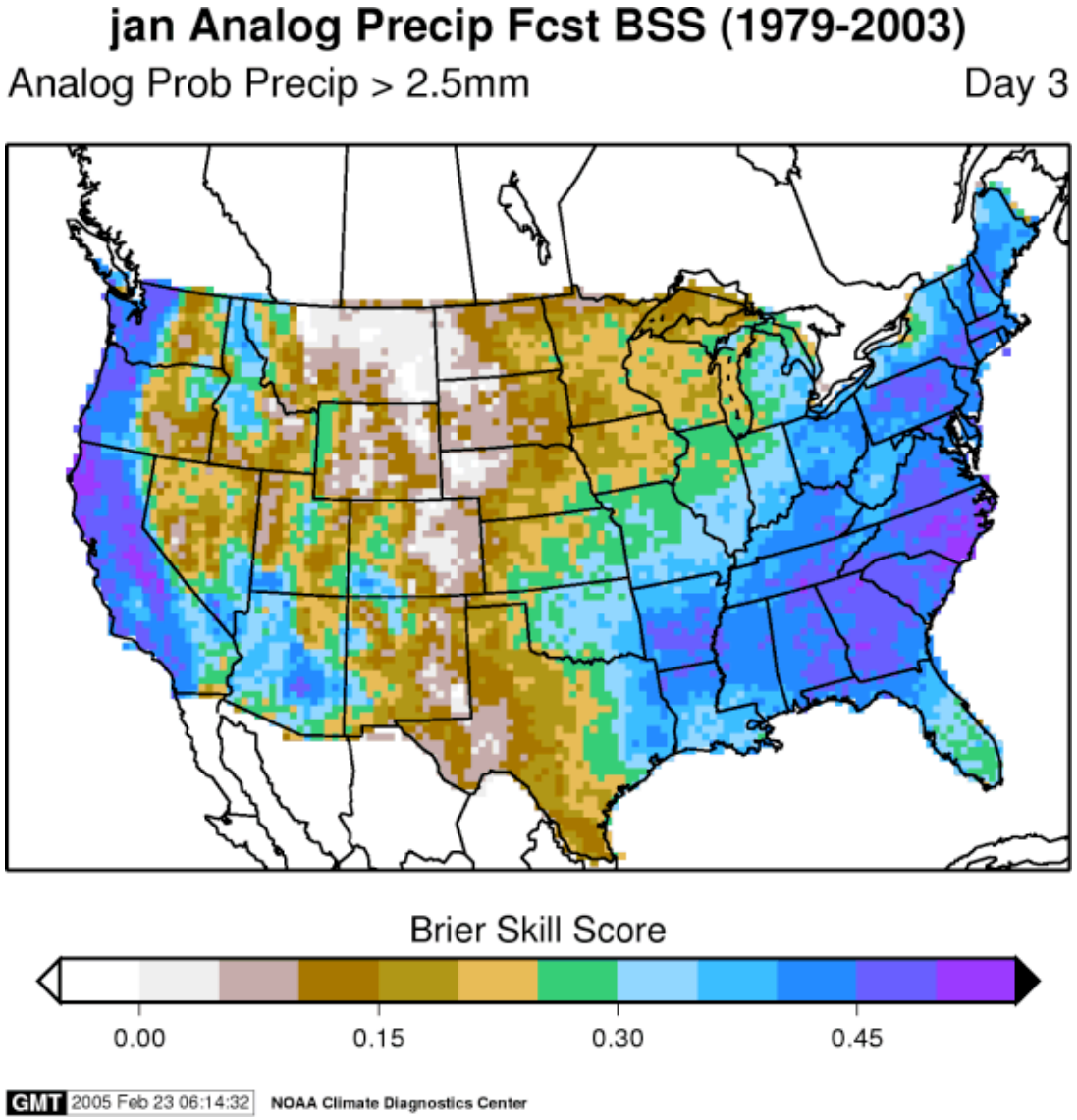
Reforecast Analogs

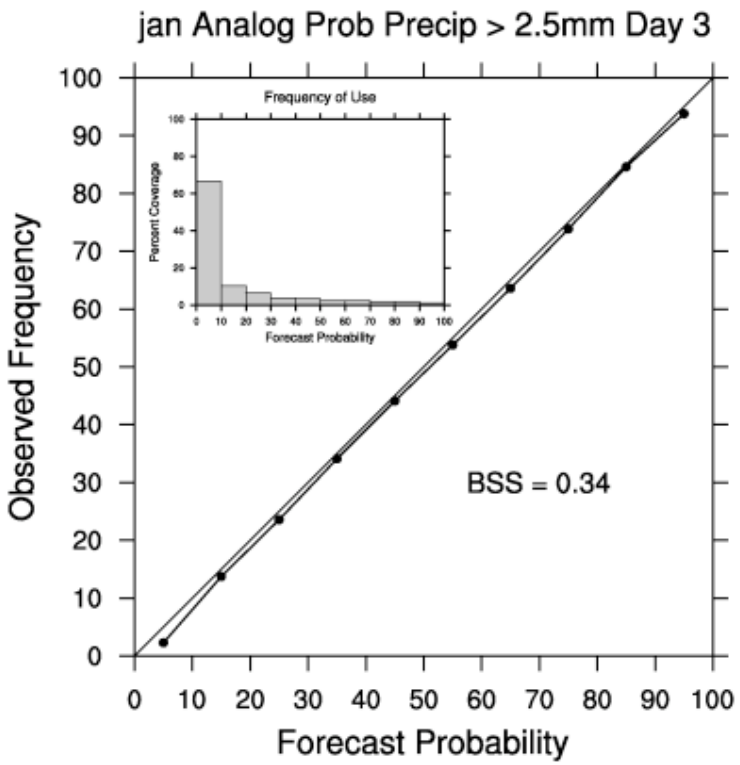


Operational GFS



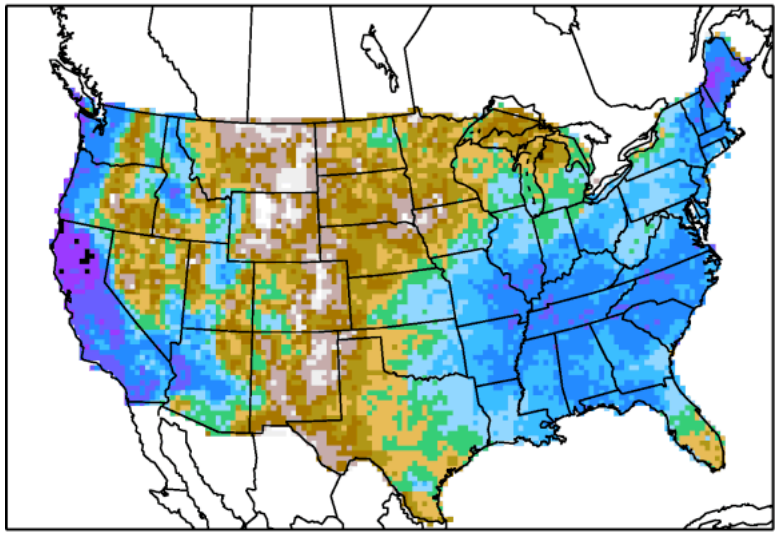
Annual Cycle of Skill





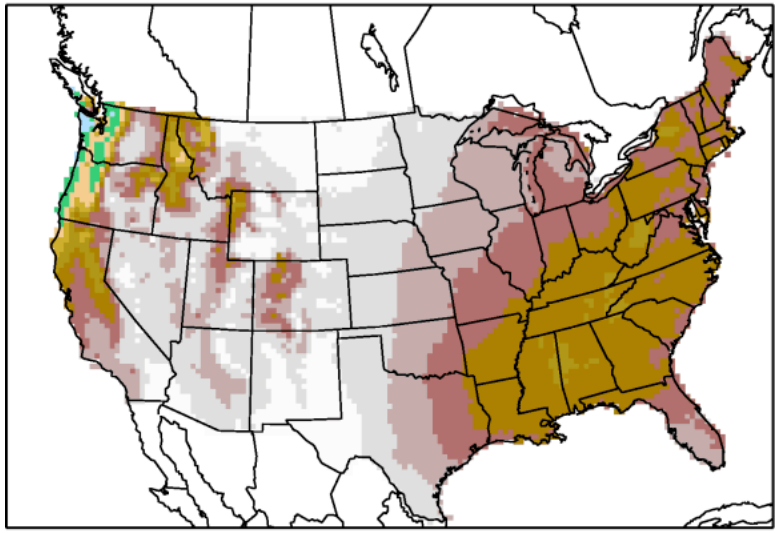
feb Analog Precip Fcst BSS (1979-2003)

Analog Prob Precip > 2.5mm Day 3



Climo Prob Precip > 2.5mm

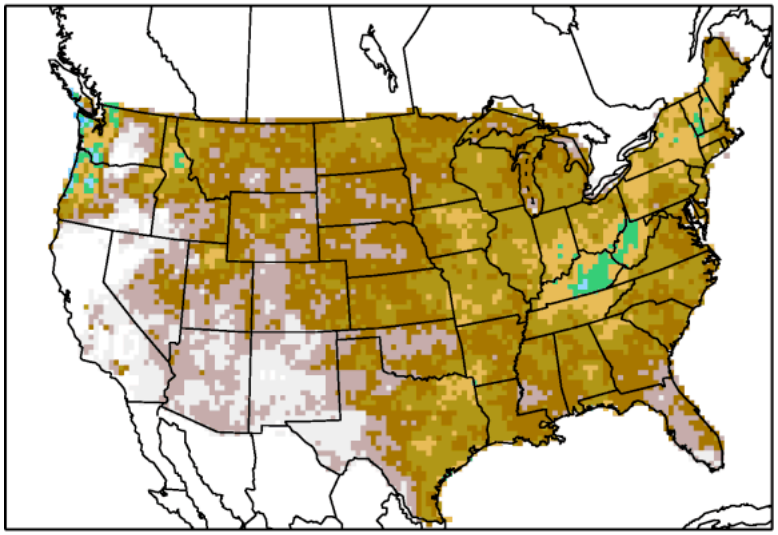
fcst from 2005022800 valid 2005030200-2005030200



GMT 2005 Feb 28 12:20:58 NOAA Climate Diagnostics Center

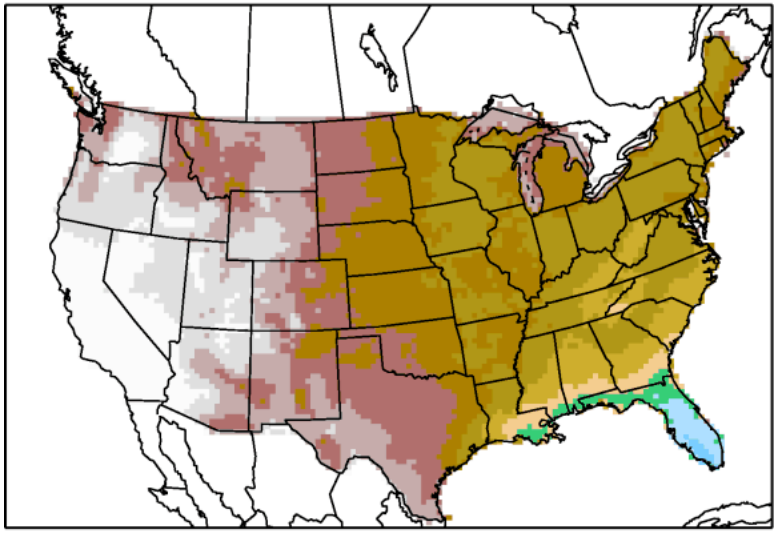
jul Analog Precip Fcst BSS (1979-2003)

Analog Prob Precip > 2.5mm Day 3



Climo Prob Precip > 2.5mm

fcst from 2004070100 valid 2004070300-2004070300

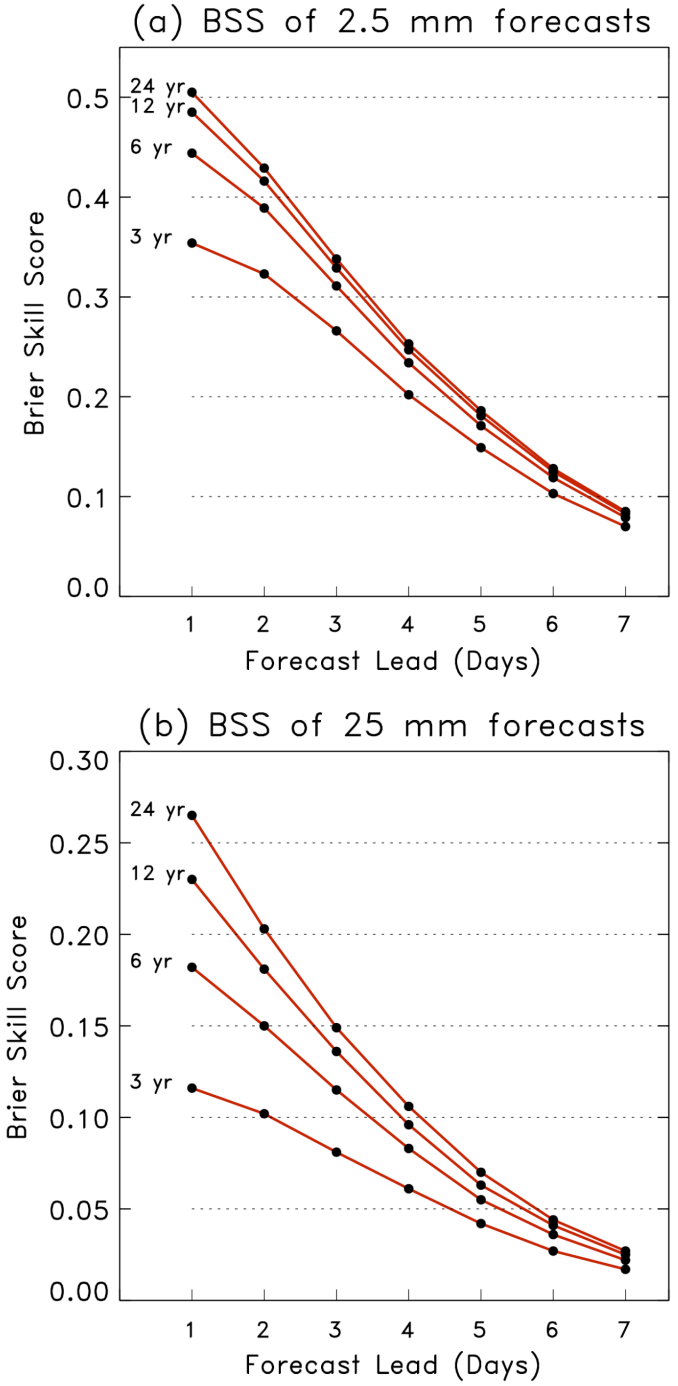


GMT 2005 Feb 28 15:59:07 NOAA Climate Diagnostics Center

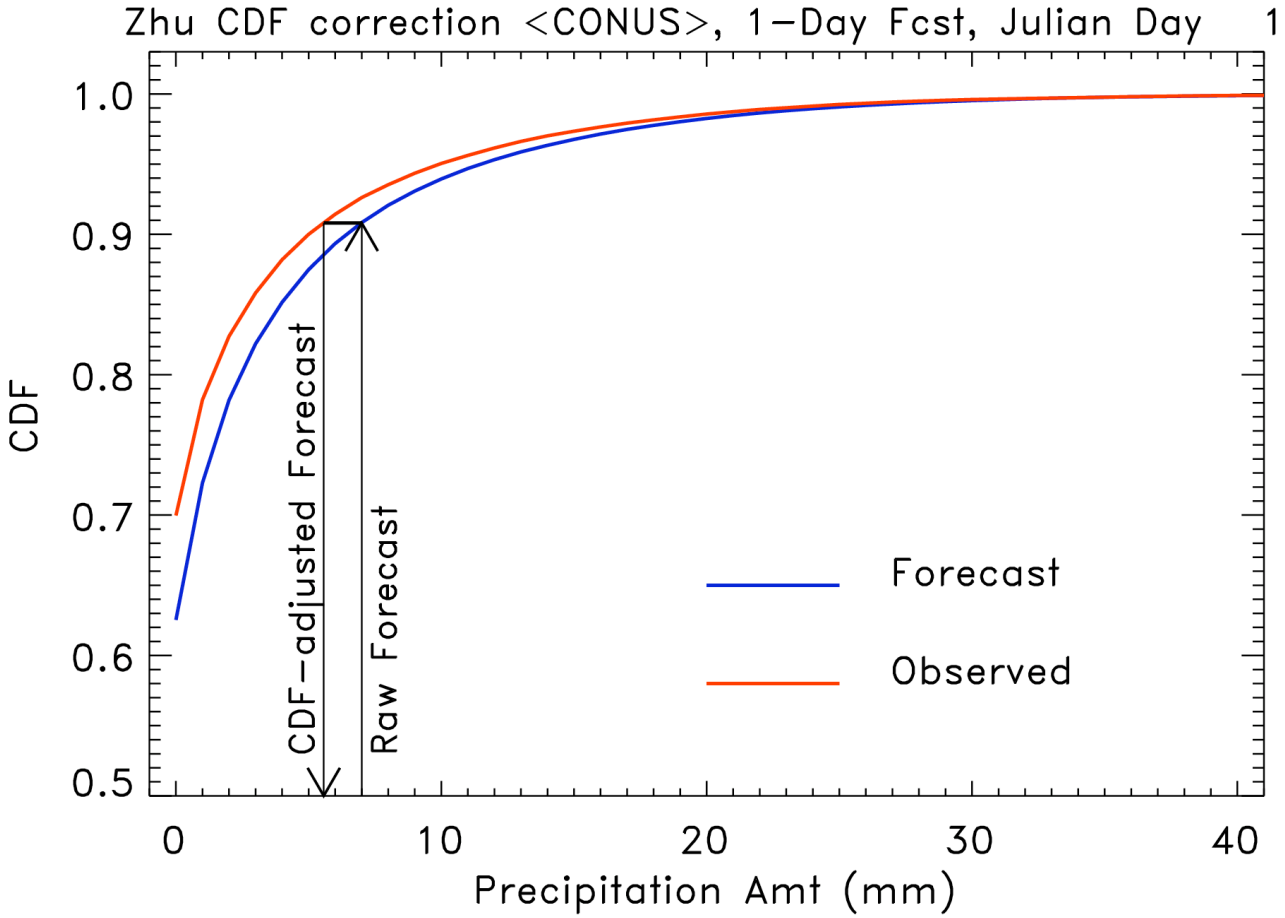
Importance of training sample size

This shows skill of precipitation forecasts the analog technique, JFM 1979-2003 data over conterminous US (CONUS).

increased sample size especially important for calibrating rarer, high-precipitation events.



Other calibration methods: NCEP technique for CDF correction (Yuejian Zhu)



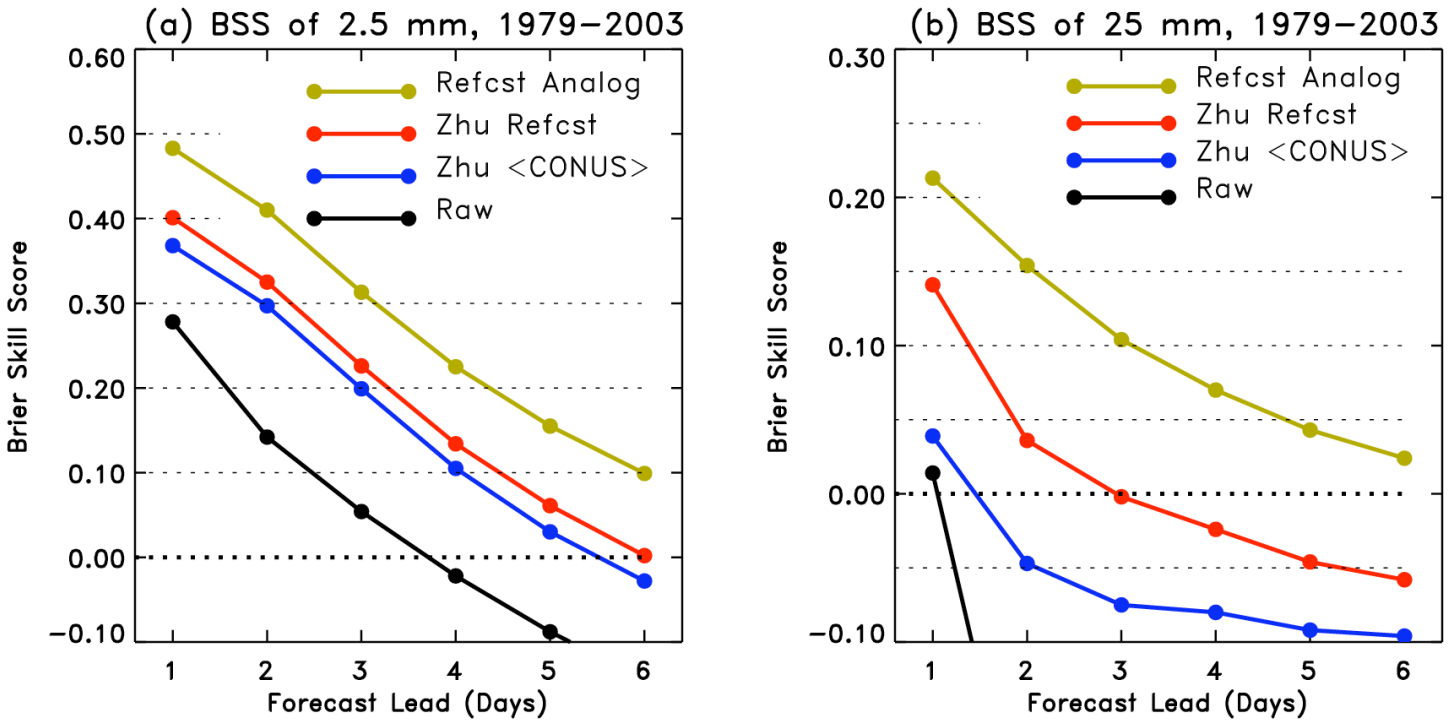
(1) Get CDFs of forecast and observed, averaged over CONUS using, say, last 30 days of data.

(2) Use difference in CDFs to correct each ensemble member's forecast. In example shown, raw 7 mm forecast corrected to ~5.6 mm forecast.

NOTE: bias only, not spread correction.

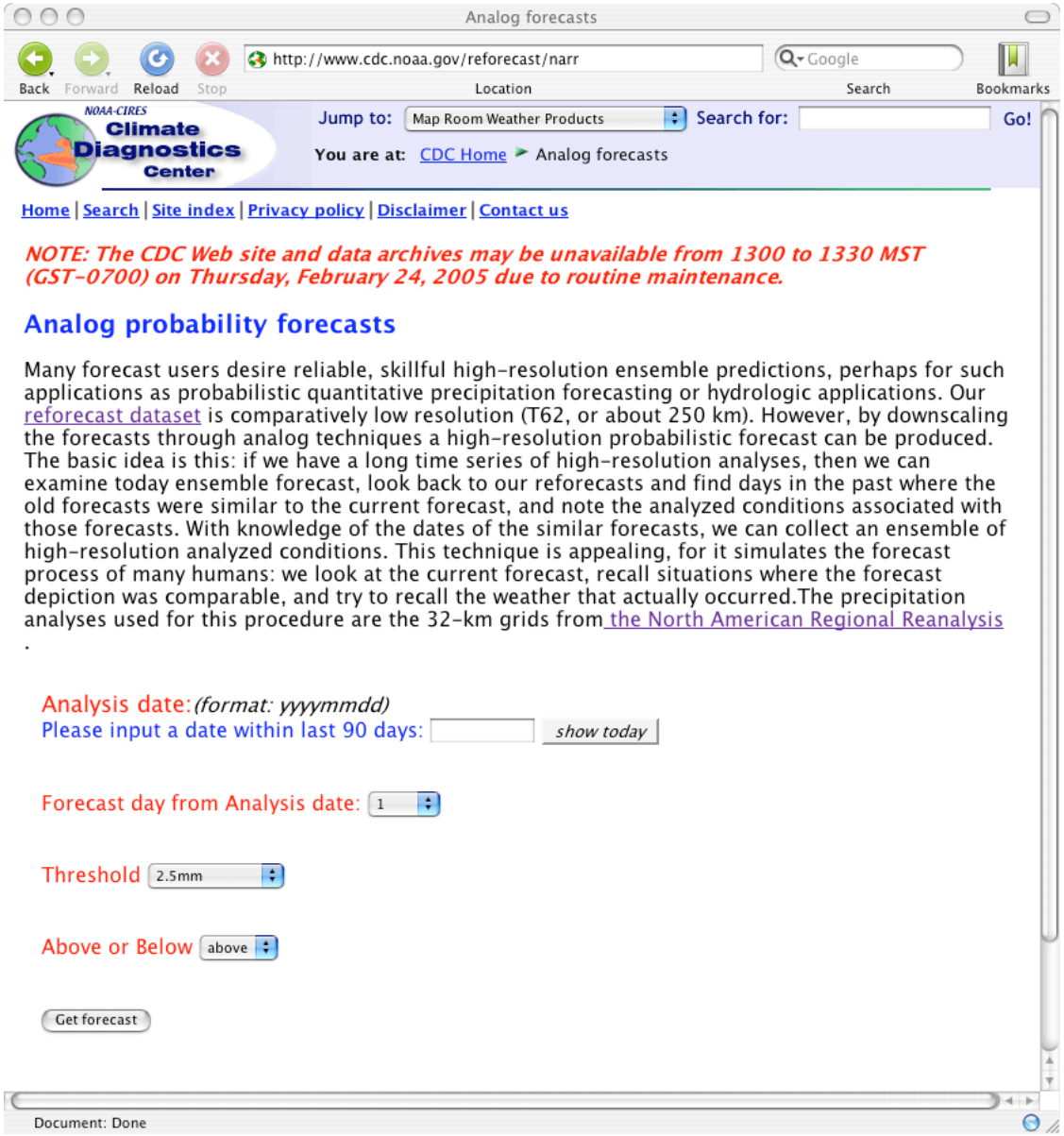
Ref: Zhu and Toth, 2005 AMS Annual Conf.

Skill of CDF correction vs analogs

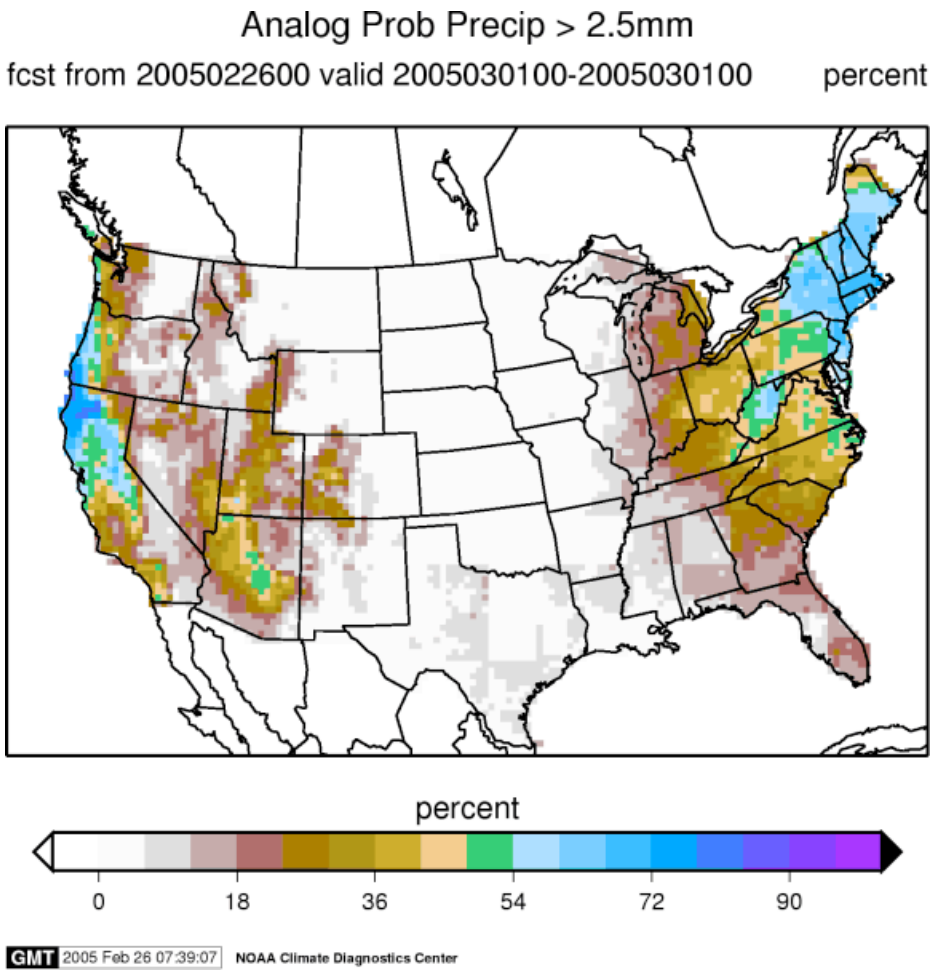


Notes: (1) Here, verification on coarse 2.5 degree grid.
(2) Zhu <CONUS> has benefit at 2.5 mm, correcting drizzle over-forecast.
(3) Location-dependent Zhu technique using reforecasts adds skill, esp. at 25 mm.
(4) Large additional skill by using analog reforecast technique, again largest at high thresholds.
(5) The type of calibration technique really matters.

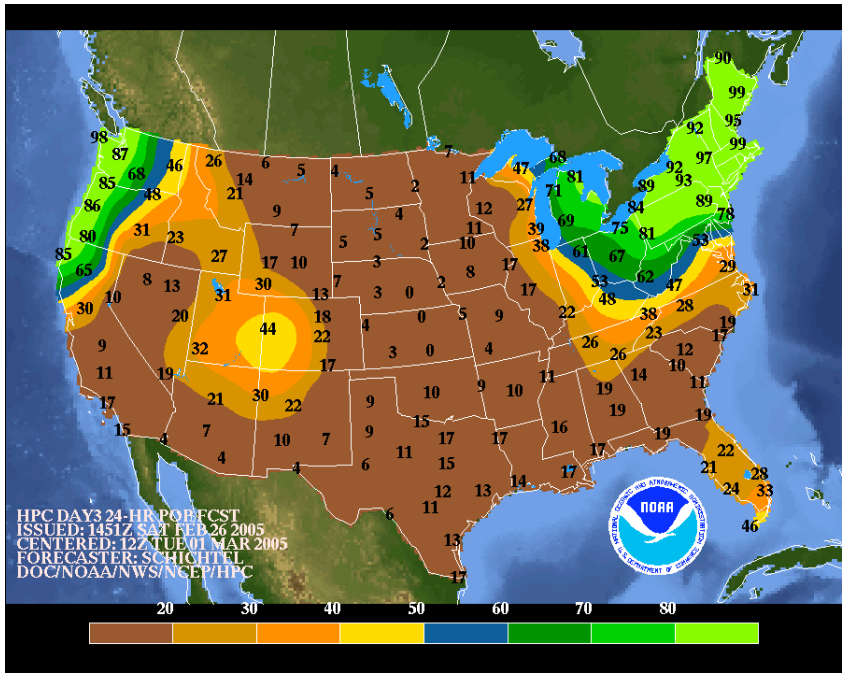
http://www.cdc.noaa.gov/reforecast/narr



Example - fcst valid yesterday

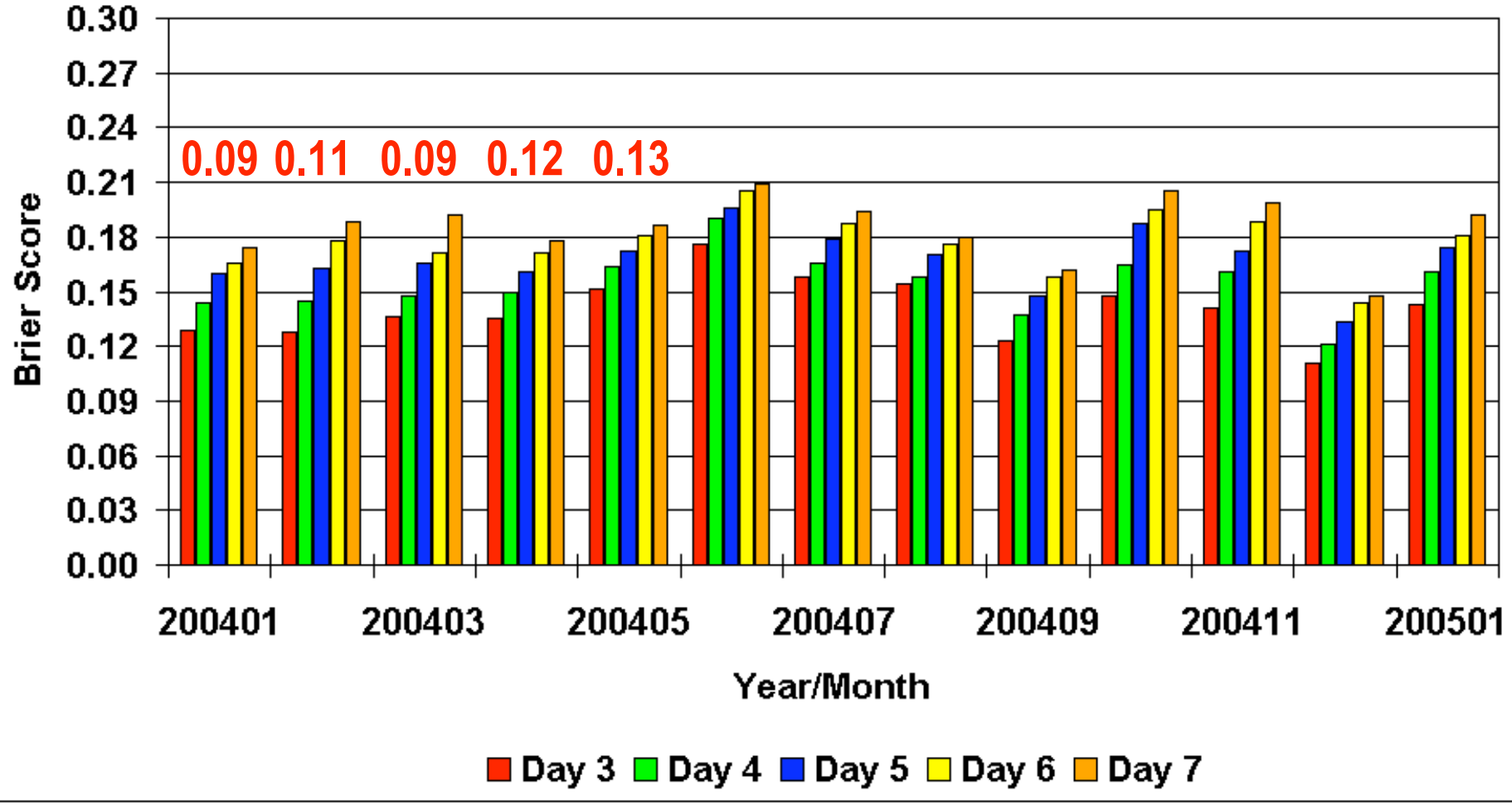


HPC 3 day POP



HPC vs REFCST POP verifications (Brier Score)

bars show HPC (days 5-7), red # shows REFCST day 5 score



Conclusions

- Big improvement in PQPF skill possible.
- Need long training dataset, especially for high amounts (rare events).
- Need good, long-term obs! (analog PQPF cannot be trusted where NARR precip is suspect).